The EMBO Journal vol.8 no.2 pp.561 – 568, 1989

Induction and repression of the *Drosophila Sgs-3* glue gene are mediated by distinct sequences in the proximal promoter

Marianne Martin, Angela Giangrande, Claude Ruiz and Geoff Richards

Laboratoire de Génétique Moléculaire des Eucaryotes du C.N.R.S., Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cedex, France

Communicated by M.Ashburner

The normal developmental expression of the *Drosophila* salivary gland secretion protein gene *Sgs-3* requires the interaction of a distal and proximal regulatory element. A deletion/replacement analysis of the proximal promoter in stably transformed lines shows that induction of an *Sgs-3/Adh* fusion gene is normal if sequences from +10 to −50 are replaced by those of the *hsp70* gene. Sequences between −98 and −50 are necessary for this expression but there is internal redundancy within this region as two distinct upstream sequences of 18 and 22 bp respectively are sufficient for stage- and tissue-specific expression, albeit at reduced levels. A point mutation at −53 eliminates the ecdysone-mediated repression of the *Sgs-3* promoter at pupariation. We report mosaicisms of expression within the salivary gland for a number of stably transformed lines.

Key words: Drosophila/salivary gland secretion protein gene

Introduction

Germ line transformation of *Drosophila* (Rubin and Spradling, 1982) has provided a powerful approach for *in vivo* studies of gene regulation in development. A number of model systems, e.g. *hsps* (Xiao and Lis, 1988), *Ddc* (Bray *et al.*, 1988) and *Adh* (Fischer and Maniatis, 1988), have now reached the point that regulatory elements have been identified with a precision of some 20 bp and are amenable to analysis with single-base-pair mutations.

In the case of the salivary gland secretion protein gene, *Sgs-3*, two key regions have been identified (see Meyerowitz *et al.*, 1987 for review). Transformants carrying sequences extending to −130 bp from the transcription start site express *Sgs-3* at low levels (Bourouis and Richards, 1985a; Vijay Raghavan *et al.*, 1986). The expression of these minimal constructs is restricted, as for the endogenous *Sgs-3* gene, to the third-larval-instar salivary glands. The second region is centred at −600 bp where a 63-bp deletion reduces activity to ~5% of wild-type levels (Ramain *et al.*, 1988). Other elements having modest (2- to 3-fold) effects on transcript levels have been identified further upstream (Giangrande *et al.*, 1987), one of which is included in the transcribed region of *Sgs-7*, the 5′ neighbour of *Sgs-3* in the 68C glue gene cluster (Garfinkel *et al.*, 1983). This organization appears to differ from that of *Sgs-4* where a remote 420-bp fragment has been shown to act as a stage- and tissue-specific enhancer when placed adjacent to the *Drosophila Adh* gene (Shermoen *et al.*, 1987). A recent analysis confirms that for *Sgs-4* tissue specificity is determined in this remote region and suggests that there is functional redundancy within this fragment (Jongens *et al.*, 1988).

The *Sgs-3* gene has an intriguing hormonal regulation. *In vivo* and *in vitro* experiments with wild-type and mutant strains suggest that *Sgs-3* is both induced, in the mid third larval instar and later, repressed, prior to pupariation, by the moulting hormone ecdysone [see Meyerowitz *et al.* (1987) for references]. As the developmental expression of the minimal constructs (see above) is normal, we should expect to find hormonally responsive element(s) within the 130-bp fragment.

For our present studies we have chosen to work with a fusion gene. Such constructs have been successfully exploited for developmental studies of other *Drosophila* promoters, e.g. *YP1* (Garabedian *et al.*, 1986), *Sgs-4* (Shermoen *et al.*, 1987), *hsps* (Xiao and Lis, 1988), *ftz* (Hiromi *et al.*, 1985), *Adh* (Fischer and Maniatis, 1988), chorion genes (Romano *et al.*, 1988) and *Ubx* (Bienz *et al.*, 1988), and stage- and tissue-specific expression has been observed with *Sgs-3* fusion gene constructs, *Sgs-3/Adh* (Bourouis and Richards, 1985b) and *Sgs-3/lacZ* (Vijay Raghavan *et al.*, 1986). This approach should enable us to detect any aberrant regulation, e.g. novel stage or tissue expression, resulting from our *in vitro* mutations within the proximal 130-bp fragment.

In this study we have used sets of overlapping oligonucleotides to synthesize deletion and replacement derivatives of the proximal 130-bp fragment within our reference construct which carries 2.7 kb of *Sgs-3* 5′ sequence. We show that *Sgs-3*-specific expression is retained when the TATA-box-containing region is replaced by corresponding sequences from a heterologous promoter (*hsp70*). We present evidence that the proximal upstream sequences may include a number of elements sufficient for *Sgs-3*-cell- and stage-specific expression, but require the presence of the distal upstream sequences for efficient expression. Finally a single-base-pair mutation shows that different *cis*-acting DNA sequences mediate *Sgs-3* induction and repression.

Results

Analysis of Sgs3/Adh fusion gene products

We characterized the developmental expression of the *Sgs-3/Adh* fusion gene glAd (Bourouis and Richards, 1985b) transformed in the Adh^{n248}, ry^{506} strain in several independent lines (Figure 1 and Table I) by histochemical staining of larval, pupal and adult stages. ADH activity was restricted to third larval instar and white prepupal salivary glands (Figure 2a, b and c). During the third larval instar ADH activity was first observed in the most posterior cells of the gland and spread towards the anterior so that at the end of the instar all but the most proximal cells were stained. This parallels the accumulation of salivary gland secretion proteins as detected by periodic acid Schiff reagent (PAS) staining

(see Berendes and Ashburner, 1978). Transcription of *Sgs-3* was repressed before pupariation when the glue proteins were secreted. In contrast the fusion gene product was detected within the salivary gland cells of white prepupae (Figure 2c). This resulted both from the fact that the signal peptide sequence of *Sgs-3* was not present in the fusion and the stability of the ADH protein (see, for example, Bonner *et al.*, 1984).

We analysed transcripts of the fusion gene and the endogenous *Sgs-3* gene in third instar larvae by primer extension analysis using a mixture of two oligonucleotide primers. The *Sgs-3* primer gives rise to a major 54-bp elongation product. For the fusion gene we used either of two primers, A and B (see Materials and methods). Primer A gives rise to a 53/54-bp doublet for the larval *Adh* transcript of the Adh^{n248} strain and a 60-bp product for the fusion gene (Figure 3a). Primer B gives rise to a 35/36-bp doublet for the larval *Adh* transcript and a 42-bp product for the fusion gene (Figure 3b). In both cases there was a minor product corresponding to initiation at −1 for *Sgs-3* as well as the fusion gene (small arrows, Figure 1). The *Adh* primer B eliminated the problem of coincidence of one larval *Adh* product with that of the endogenous *Sgs-3* transcript. There was a normal developmental expression of *Adh* larval transcripts in Adh^{n248} which was nonetheless an ADH null strain because of an internal duplication in the coding region (Chia *et al.*, 1985). These *Adh* larval transcripts were negligible in late postfeeding larvae but in younger larvae may lead to an overestimation of the expression of *Sgs-3* (see Figure 3, glAd:7 lanes). Unfortunately primer B proved less stable than primer A so that the ratio fusion gene:*Sgs-3* transcripts decreased dramatically with time when aliquots from the same RNA sample were run as internal standards on successive gels (not shown). When using the *Sgs-3* primer and the *Adh* primer A labelled to a comparable specific activity, it appears that the transcripts of the fusion gene were

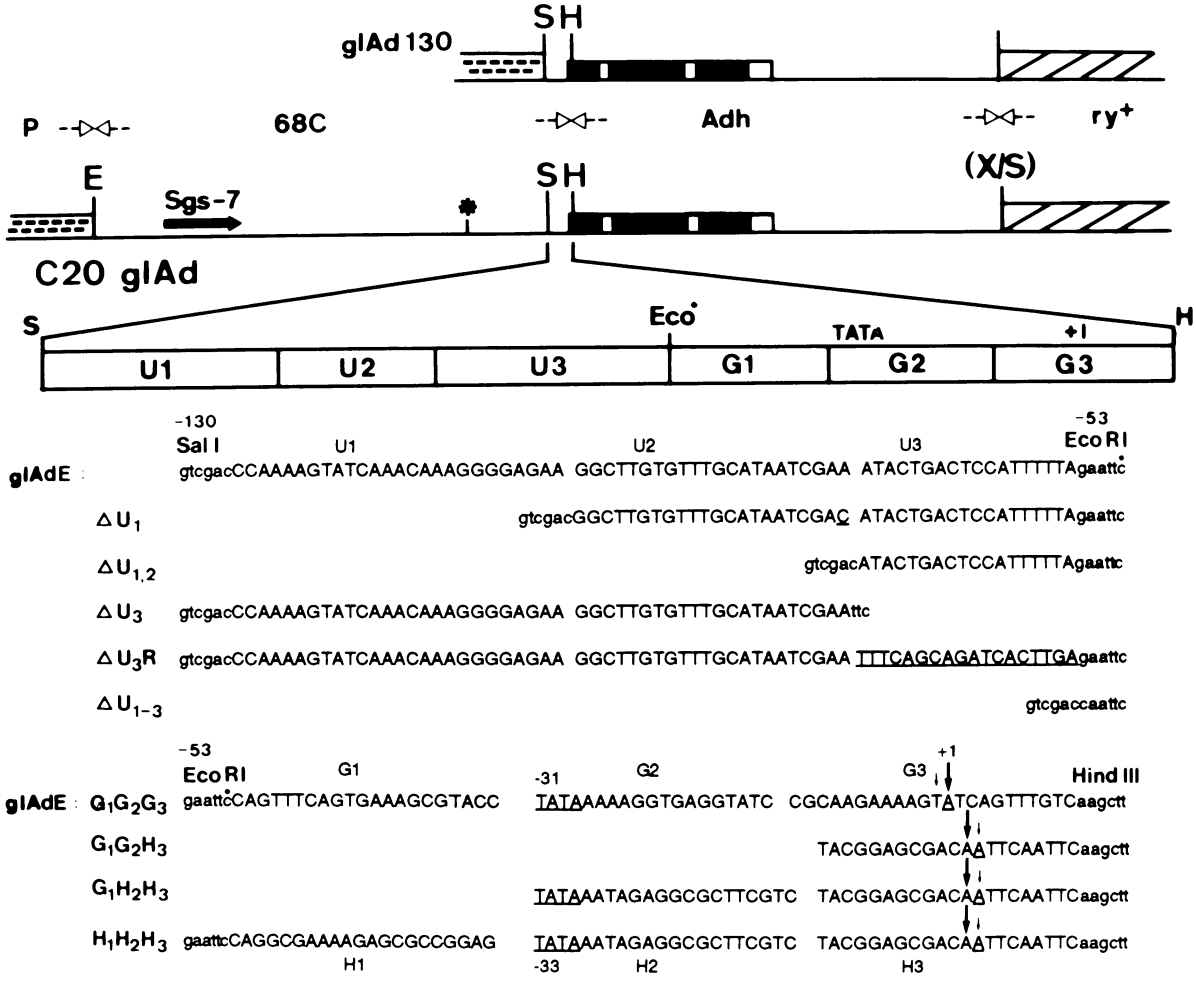

Fig. 1. *Sgs-3/Adh* fusion gene constructs. The reference construct C20glAd (Bourouis and Richards, 1985b) is shown centrally with the source of the DNA fragments: P element sequences (dotted box segment), 68C DNA from the *Eco*RI site at −2.7 kb from the *Sgs-3* startsite to the *Hin*dIII linker at +10 on the *Sgs-3* leader, the Adh fragment containing the larval exon (solid box) and intron (open box) sequences and the Xdh ry^+ fragment (hatched box) that serves as a marker gene in Carnegie 20 (Rubin and Spradling, 1983). In the 68C fragment the neighbouring *Sgs-7* transcript is denoted by an arrow. An asterisk at −600 from *Sgs-3* denotes the distal regulatory element discussed in the text. The minimal construct glAd130 is shown above. In glAd130I the fusion gene is inverted with respect to the *P* and ry^+ fragments (not shown). The *Sal*I−*Hin*dIII fragment is expanded below and the *Eco*RI site created by site-directed mutagenesis is marked by an asterisk. The G→C point mutation is marked by a solid point on the DNA sequence of the resulting glAdE shown below, which is divided into the upstream regions U_1-U_3 and TATA box regions G_1-G_3. The derivatives assembled from overlapping sets of oligonucleotides are shown below, the flanking restriction sites being given in small letters, except for those nucleotides considered to form part of a box. In ΔU_1 the point mutation introduced at −77 is underlined (see text). In ΔU_3R the replacement sequence, introduced so as to conserve the spacing of flanking regions, is underlined. In the G_1-G_3 replacement series, as an aid to clarity, the G_1 and G_2 boxes are shown only for glAdE. The TATA sequence and the reported transcription startsites (see text) are underlined. Large and small arrows represent the major and minor startsites observed with these constructs.

Table I. Transformed lines and fusion gene expression

glAd	:5 (3)	+++	glAdE-$G_1G_2G_3$	:1 (3)	++
	:6 (2)	+++		:2 (X)	+++
	:7 (2)	+++		:3 (3)[a]	+++*
	:8 (3)	+++			
			$G_1G_2H_3$	:1 (3)	++
glAd130	:1 (2)	−		:2 (2)	+++
	:2 (X)	−		:3 (2)	+++
	:3 (2)	−		:4 (X)	+++
	:4 (3)	−		:5 (2)[a]	+++
	:5 (3)[a]	−			
	:6 (3)	−	$G_1H_2H_3$	:1 (2)	+++*
	:7 (2)	−		:2 (3)	++
	:8 (X)	−		:3 (3)[a]	++
				:4 (X)	+++
glAd130I	:1 (2)[a]	−		:5 (3)	++
	:2 (3)[a]	−		:6 (X)	++
	:3 (2)	−			
	:4 (3)	−	$H_1H_2H_3$	:1 (X)	+++*
				:2 (X)	++*
				:3 (2)	+++*
				:4 (X)	++*
				:5 (X)[b]	+++

After the designation of each line the chromosome carrying the insertion is given in parentheses: [a]homozygous lethal insertion, [b]a male lethal—analysis performed on female larvae.
+++, heavy, ++, intermediate, +, light staining, − no staining.
An asterisk denotes lines where individuals showed patchy staining (see Figure 2 and text).

at levels similar to those of the resident *Sgs-3* gene (see Figure 3a for examples).

Vijay Raghavan *et al.* (1986) showed that β-galactosidase activity from an *Sgs-3/lacZ* fusion gene containing 130 bp of *Sgs-3* 5′ sequence was detected specifically in third instar salivary glands. We constructed a derivative of our *Sgs-3/Adh* fusion gene, that retained the same 130 bp of 5′ sequence, inserted in both orientations in the polylinker of Carnegie 20 (glAd130 and 130I, Figure 1) and obtained several transformed lines for each construction (Table I). We were unable to detect activity of the fusion gene either by histochemical staining or by primer extension analysis of transcripts (data not shown—see Discussion) and therefore chose to modify the 130-bp region within the glAd transposon, thus retaining the remote 5′ elements necessary for abundant expression (see Introduction).

Experimental strategy

In order to dissect the proximal regulatory sequences we introduced a point mutation (G→C) at position −53 thus creating an *Eco*RI site. This arbitrarily divides the region into two subregions from *Sal*I to *Eco*RI and *Eco*RI to *Hin*dIII in the fusion gene leader sequence. Transformed lines carrying the resulting glAdE (see Figure 1 and Table I) were analysed both by staining for ADH activity in larval salivary glands and by the primer extension assay as above. The introduction of the *Eco*RI site had no apparent effect upon the induction of fusion gene activity in larval salivary glands

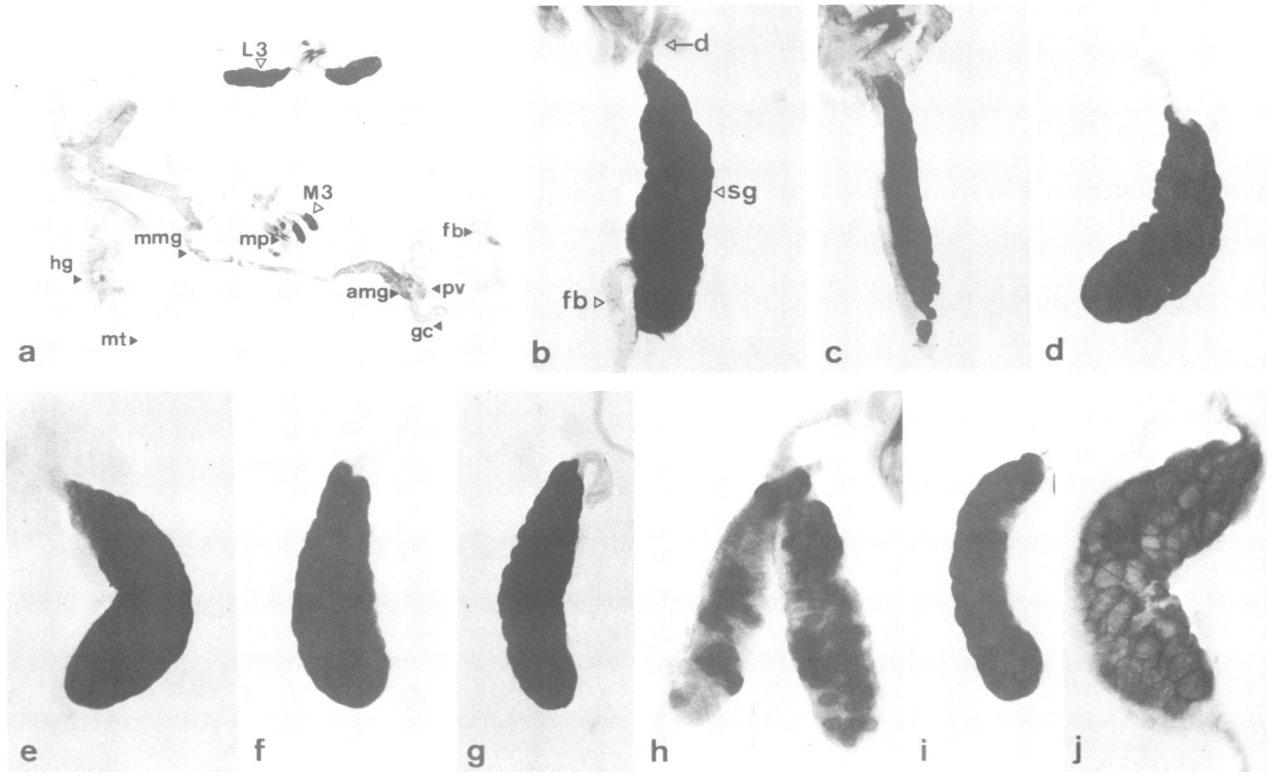

Fig. 2. Histochemical staining of ADH activity in *b* Adh^{n248}; ry^{506} strains transformed with the *Sgs-3/Adh* fusion gene derivatives. (**a**) General view of tissues from a mid third instar larvae from the line glAd:6 stained for ADH, together with a mouthparts/salivary gland complex of a late third instar larva. amg, anterior mid gut; mmg, middle midgut; hg, hindgut; fb, fat body; gc, gastric caeca; mp, mouthparts; mt, Malpighian tubules; pv, proventriculus. Salivary glands are denoted M3 and L3 for the mid and late third instar larva respectively. (**b-j**) Late third instar (except **c**) salivary glands of glAd:6 (**b**), glAd:8, white prepupae (note that ADH activity remains in the collapsed gland after glue expulsion) (**c**), glAdE:2 (**d**), $H_1H_2H_3$:5 (**e**), ΔU_1:5 (**f**), $\Delta U_{1,2}$:2 (**g**), $\Delta U_{1,2}$:4 (**h**), ΔU_1H_{1-3}:2 (**i**) and $\Delta U_1\Delta U_3RH_{1-3}$:3 (**j**). In (**b**) the salivary gland (sg) its duct cells (d) and attached fat body cells (fb) are indicated.

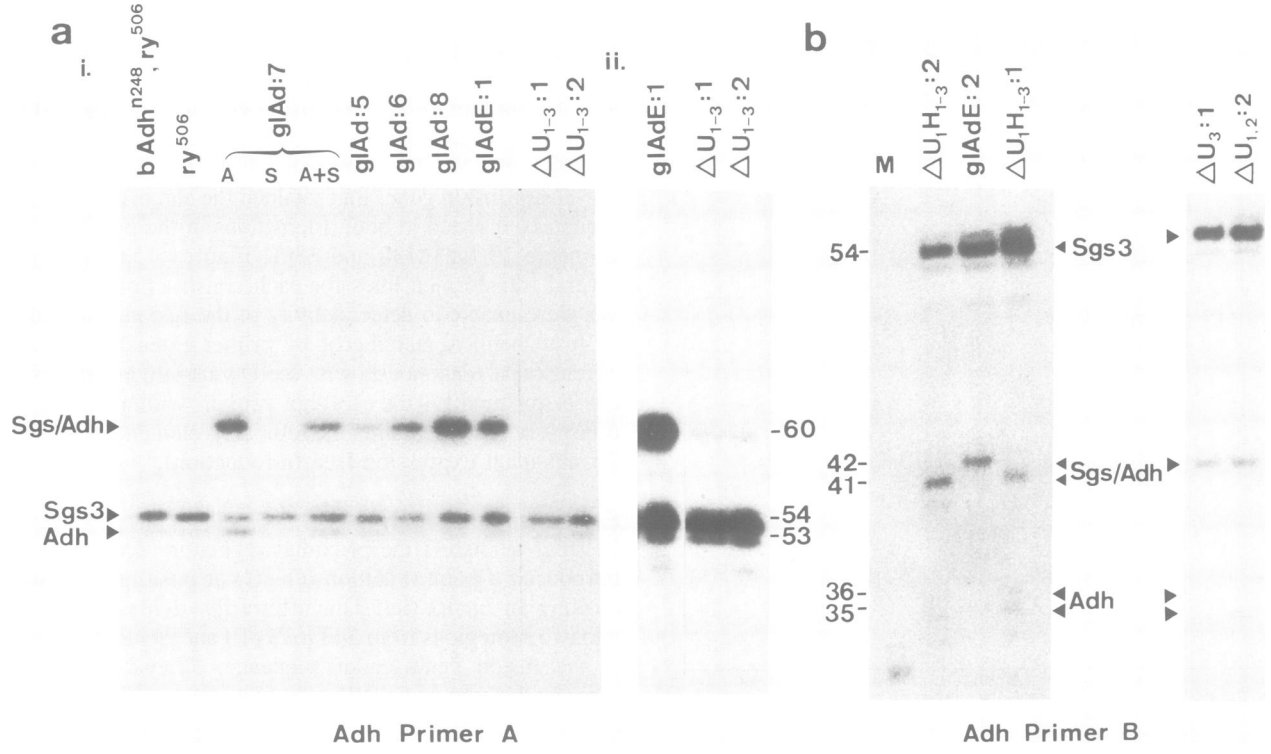

Fig. 3. Primer extension analyses of endogenous *Sgs-3* gene and fusion gene transcripts in third instar larvae. Five micrograms of total RNA were hybridized with the *Sgs-3* and/or the *Adh* (A or B) primer as described in Materials and methods. The extension products from the endogenous *Sgs-3* gene (Sgs3), the larval *Adh* transcription unit (Adh) and the transformed *Sgs-3*/*Adh* gene (Sgs/Adh) are indicated by arrowheads. Transformed lines are denoted as in Tables I and II. In (**a**)**i** the products from non-injected host lines *b Adh*^{n248}, *ry*^{506} and *ry*^{506} are shown first. For glAd:7 three aliquots were hybridized with either the *Adh* primer A (A), the *Sgs-3* primer (S) or both (A+S). All other lines are hybridized with both primers. In (**a**)**ii** the last three lanes of **i** are shown following extended exposure of the autoradiograph so as to reveal the minor band in ΔU_{1-3} lanes. In (**b**) samples are analysed with the *Sgs-3* and *Adh* primer B. M denotes a marker lane.

(Figures 2d and 3a). We then divided each subregion arbitrarily into three boxes U_1-U_3 and G_1-G_3 (see Figure 1, bottom). For the upstream sequences we constructed deletion and substitution derivatives, whilst for the TATA box and startsite region we replaced *Sgs-3* sequences (G_1-G_3) with the corresponding sequences (H_1-H_3) from the *hsp70* promoter. The *hsp70* sequences extend to −50 (thus avoiding the inclusion of a heat shock element), include a consensus TATA box and startsite and have been used previously as a 'neutral' promoter in *Drosophila* fusion gene studies (see Discussion for examples).

Sgs3/hsp70 promoter fusions

To determine whether tissue or developmental specificity are a property of the sequences surrounding and including the *Sgs-3* TATA box, we constructed, by shotgun ligation of oligonucleotides (Grundström *et al.*, 1985), transposons in which *Sgs-3* sequences (or $G_1G_2G_3$) were progressively replaced by corresponding sequences from the *hsp70* promoter to give glAdE-$G_1G_2H_3$, $G_1H_2H_3$ and $H_1H_2H_3$. In the first, the reported *hsp70* startsite (see Hultmark *et al.*, 1986, underlined Figure 1) was at 31 bp from the *Sgs-3* TATA, thus respecting the distance found in *Sgs-3*. In $G_1H_2H_3$ and $H_1H_2H_3$ the *Sgs-3* TATA motif was replaced by that of *hsp70* and the startsite was 33 bp downstream, as in *hsp70*. In all three the leader sequence was shortened by 2 bp. Histochemical staining of the transformed lines (Table I, Figure 2e for example) revealed a normal developmental expression, restricted to third instar and white prepupal salivary glands (see above). Primer extension analysis (with primer B) of $G_1G_2H_3$, $G_1H_2H_3$ (Figure 4) and $H_1H_2H_3$ (not shown) revealed a major product of 41 bp and a minor product of 40 bp indicating that initiation occurs primarily at 30 bp from the *Sgs-3* TATA for the former and at 32 bp from the *hsp70* TATA for the others (see arrows, Figure 1). These assignations were confirmed by direct sequencing of the transcripts (data not shown) and are in agreement with the *hsp70* startsite proposed by Hultmark *et al.* (1986) from their analysis of conserved sequences of *hsp* genes. Note here that these substitutions do not significantly reduce transcript levels (see Figure 4). These results suggest that the regulatory elements determining the stage- and tissue-specific induction of *Sgs-3* are not contained in the region −50 to +10.

The proximal upstream sequences

In parallel experiments we investigated the *SalI*−*Eco*RI fragment by a series of deletion and replacement constructs, again based on a series of three boxes (U_1-U_3) (see Figure 1 and Table II for the transformed lines). When all three were deleted (ΔU_{1-3}) the ADH histochemical assay revealed that the large majority of larvae were ADH negative whilst there was weak staining of a few salivary gland cells in ~1 in 10 larvae. Primer extension analysis of RNA extracted from a pool of 30 larvae revealed a minor band upon extended exposure of the autoradiograph (Figure 3a, i and ii) whilst analyses of individual larvae confirmed that expression was limited to exceptional larvae (data not shown). Taken together these results show that this region contains an element(s) indispensable for a normal expression

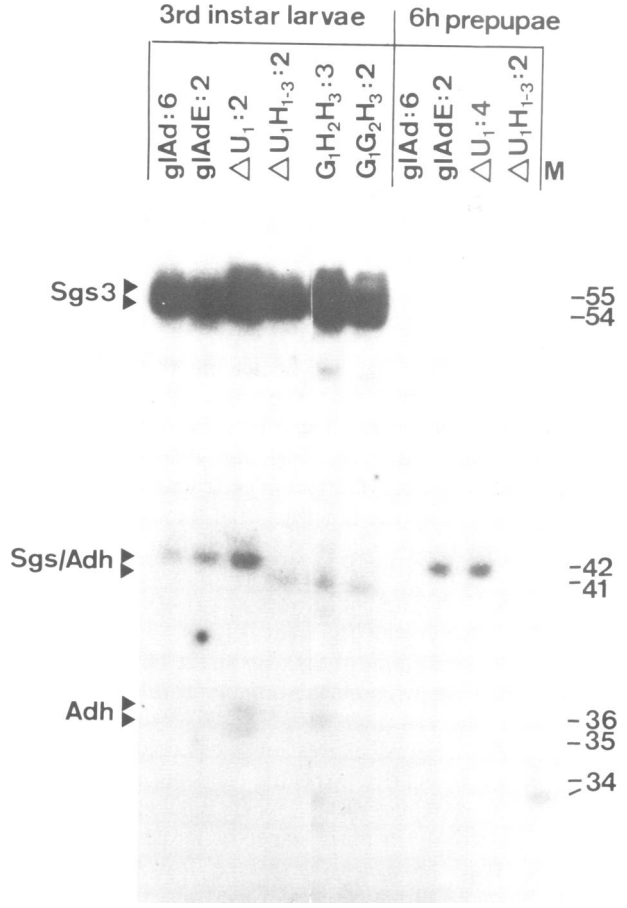

Fig. 4. Primer extension analyses of endogenous *Sgs-3* gene and fusion gene transcripts in late third instar larvae and 6-h prepupae. Five micrograms of total RNA were hybridized with the *Sgs-3* and *Adh* B primers, and treated as described in Materials and methods. The extension products from the endogenous *Sgs-3* gene (Sgs3), the larval *Adh* transcription unit (Adh) and the transformed *Sgs-3/Adh* gene (Sgs/Adh) are indicated by arrowheads. The overexposed *Sgs-3* region includes both the major, 54-, and minor, 55-base extension product. Sizes (in bases) are given at the right of the figure next to the marker lane (M). Transformed lines are denoted as in Tables I and II. The extension products of H_3 carrying fusions are one base shorter than those carrying G_3 (see Figure 1 and text). Note the absence of endogenous *Sgs-3* and larval *Adh* transcripts in 6-h prepupae.

of the fusion gene.

The deletions ΔU_1 and $\Delta U_{1,2}$ removed sequences from the *SalI* site (−130) to −98 and −81 respectively. Both carried a point mutation A→C at −77, introduced so as to reconstitute a *SalI* site in the latter. This point mutation was without effect, as in the deletion construct ΔU_1 there was uniform ADH staining of glands that was indistinguishable from that of glAd transformants in all lines (Figure 2f) and larval transcript levels were comparable to those of the glAd construct (Figure 4). Note also that this change occurs naturally in *Sgs-3* genes of related species (Martin *et al.*, 1988). When both U_1 and U_2 were deleted, $\Delta U_{1,2}$, a greater heterogeneity in expression was observed between transformed lines (Figure 2g and h). Three lines showed strong uniform colouration, one line uniform but weak staining, whilst two showed patchy expression, i.e. groups of cells within the salivary gland showed ADH activity whilst the remainder were ADH negative (Table II). Primer extension analysis of lines showing strong colouration revealed

Table II. Transformed lines carrying modified upstream sequences and fusion gene expression

ΔU_1	:1 (2)	+++	ΔU_3R	:1 (2)	++*
	:2 (3)	+++		:2 (2)	++*
	:3 (X)	+++		:3 (2)	+*
	:4 (3)	+++		:4 (X)	++
	:5 (2)	+++		:5 (2)	++
	:6 (2)	+++		:6 (X)	+++
	:7 (2)	+++			
$\Delta U_{1,2}$	:1 (3)	++*	ΔU_1H_{1-3}	:1 (3)[a]	++*
	:2 (2)	+++		:2 (3)	+++
	:3 (3)	+++		:3 (X)	+++
	:4 (3)	+*		:4 (3)	++*
	:5 (X)	++		:5 (3)	++*
	:6 (3)	+++*			
ΔU_{1-3}	:1 (3)	−*	$\Delta U_1\Delta U_3RH_{1-3}$	:1 (2)	+++*
	:2 (X)	−*		:2 (2)	+++
				:3 (2)	+++*
ΔU_3	:1 (2)	+++		:4 (3)	+
	:2 (X)	++			
	:3 (X)	+*			
	:4 (2)	++*			
	:5 (X)	++*			
	:6 (2)[a]	+++			

All symbols as for Table I, except −* which refers to the ~1 larva in 10 of ΔU_{1-3} lines which showed staining in a few gland cells.

transcripts at levels comparable to those of glAdE lines (Figure 3b).

These results suggested that the presence of the region U_3 alone is sufficient for cell-specific expression. Whether this region is indispensable or not was tested with the constructs ΔU_3 and ΔU_3R (Figure 1), the latter containing neutral DNA sequences which restore the normal spacing of the flanking sequences. In both cases ADH activity in the transformed lines (Table II) was similar to that seen with $\Delta U_{1,2}$, i.e. some lines stained uniformly whilst others showed patchy expression. Equally, transcript levels in ΔU_3 lines showing strong colouration were comparable to those of $\Delta U_{1,2}$ and glAdE lines (Figure 3b).

This analysis suggests that either both regions U_2 and U_3 contain sequences sufficient for cell-specific expression (we eliminate U_1 as its deletion has no apparent effect—see above) or that a key element is contained in the 4-bp *TaqI* site that is retained in the $\Delta U_{1,2}$, ΔU_3 and ΔU_3R constructs (see Figure 1). It is unlikely that the element consists solely of the *TaqI* sequence (TCGA) as this is present equally in the ΔU_{1-3} construct. If the former is true, i.e. there are several sequences involved, the effect of each in isolation is not as efficient as their combined effect, as seen by a reduction in the intensity of staining and/or the number of gland cells expressing the fusion gene in certain lines. Note also that, despite the insertion of a 21-bp fragment, expression in ΔU_3 and ΔU_3R lines is similar, suggesting that the position of the upstream elements relative to the promoter is not critical. Also, somewhat surprisingly, given that all contructs retain 2.6 kb of unchanged flanking sequence, these deletion constructs are sensitive to position effects in that phenotypes differ between different insertions of the same construct.

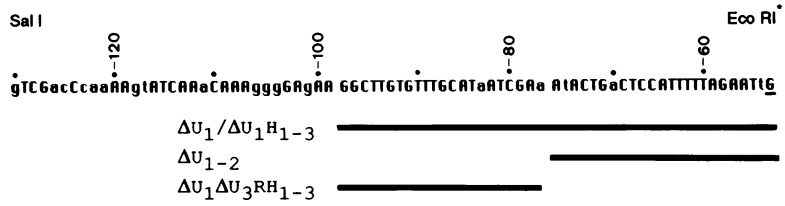

Fig. 5. A comparison of the conserved and functionally tested sequences in the *D.melanogaster Sgs-3* region −130 to −53. Nucleotides for which variants have been found in three related *Drosophila* species (Martin *et al.*, 1988) are shown in lower-case letters. The nucleotide mutated to form an *Eco*RI site (see text) is underlined. The fragments from this region included in our minimal constructs that retain tissue and stage specificity are shown schematically below by bar segments.

A minimal Sgs-3 proximal promoter

The above results suggest that there are several sequences in the 130-bp 5′ sequence sufficient for a cell-specific expression of *Sgs-3*. We have synthesized two derivatives which eliminate the U_1 box and substitute H_{1-3} for the G_{1-3} boxes. In ΔU_1H_{1-3} *Sgs-3* sequences from −98 to −50 are retained, whilst in $\Delta U_1\Delta U_3RH_{1-3}$ the only *Sgs-3* sequences are those from −98 to −77 together with the −58 to −50 region carrying the synthetic *Eco*RI site. For ΔU_1H_{1-3}, ADH staining is similar to that seen in glAd and glAdE lines (Figure 2 compare panels b, d and i) as are transcript levels (Figure 3b). In $\Delta U_1\Delta U_3RH_{1-3}$, ADH staining is gland specific and equally intense for individuals in at least three independent lines, although there is a variability between lines similar to that seen in all constructs where the U_2 or U_3 boxes are deleted (Table II, Figure 2j). Thus the upstream 22-bp *Sgs-3* sequence retained in this construct is sufficient for the induction of tissue- and stage-specific expression of the fusion gene.

The −53 point mutation eliminates Sgs-3 repression

The above results essentially concern the elements necessary for a normal expression of the fusion gene during the second half of the third larval instar. The presence of ADH activity in prepupal glands led us to undertake an analysis of transcripts at this stage. The endogenous *Sgs-3* transcripts are negligible or absent at pupariation (Oh or white prepupae) as are the fusion gene transcripts in glAd lines, whilst in striking contrast the fusion gene transcripts are still present in glAdE lines (data not shown). As this suggests that the gene has escaped the hormonal repression of the *Sgs-3* promoter at the end of the instar we have examined transcripts in prepupae 6 h after pupariation so as to eliminate any possible ambiguity in staging. In both glAdE and ΔU_1 lines there are abundant transcripts at this advanced stage (Figure 4), at levels comparable to those found in third instar larvae. In contrast, in ΔU_1H_{1-3} (Figure 4), $G_1G_2H_3$ and $G_1H_2H_3$ lines (not shown) we did not detect fusion gene transcripts (although all three constructs carry the *Eco*RI site) using several independent RNA preparations from these lines. We suspect that this is a consequence of the presence of the H_3 sequence (see Discussion).

Discussion

A final understanding of *Sgs-3* regulation will require a detailed knowledge of both the proximal and distal DNA elements as well as their interactions at the DNA and protein level. Until recently we have concentrated on the more remote elements (Giangrande *et al.*, 1987; Ramain *et al.*, 1988), in this report we turn our attention to the proximal region. We now show that a 66-bp deletion (between −124 and −57, ΔU_{1-3}) effectively eliminates *Sgs-3* expression although, unlike Vijay Raghavan *et al.* (1986) with their *Sgs-3/lacZ* fusion gene, we did not detect a specific expression of our *Sgs-3/Adh* fusion gene with 130 bp of 5′ flanking sequence. Note, however, that these fusion genes differ both in the reporter gene used and in the point of fusion, +948 (*Sgs-3/lacZ*) and +10 (*Sgs-3/Adh*). More recent experiments (cited in Meyerowitz *et al.*, 1987) using an *Sgs-3/Adh* fusion gene (fused at +12 of *Sgs-3*) have led them to place the 3′ limit of the proximal element at +12, although neither study eliminates a minor effect of the *Sgs-3* transcribed sequences (i.e. +10/+12 to +948) on mRNA stability. Note that *Sgs-3* expression is essentially regulated by positive stage- and tissue-specific elements and that none of the deletions or replacements we have studied to date have led to *Sgs-3* activity in other stages or tissues.

The eukaryotic consensus sequence TATA is present between 25 and 30 bp upstream of the transcription initiation site of many genes transcribed by RNA polymerase II (Breathnach and Chambon, 1981; Bucher and Trifonov, 1986, for reviews). It is as yet unclear as to whether it constitutes a specific or non-specific promoter element (see, for example, Simon *et al.*, 1988). Point mutations of the *Sgs-3* TATA sequence (C.Mettling *et al.*, in preparation) have shown that this element is essential for *Sgs-3* expression *in vivo*. In transgenic *Drosophila*, hybrid promoters using *ftz* (Hiromi and Gehring, 1987), *Adh* (Fischer and Maniatis, 1988) and *Ddc* (Bray *et al.*, 1988) proximal upstream sequences fused to *hsp70* TATA-transcription startsite fragments have led to developmental expression determined by the upstream sequences. Our results are in accord with these studies, i.e. that the *hsp70* TATA box is a 'neutral' element, and our intermediate construction ($G_1G_2H_3$) extends this concept. In the latter the *Sgs-3* TATA sequence directs the initiation of transcription on the *hsp70* startsite.

The results of our analysis of the proximal upstream sequences were initially surprising. The deletion from −130 to −98 was clearly without effect and confirmed preliminary results from the group of Meyerowitz (cited in Meyerowitz *et al.*, 1987) that the proximal regulatory element extends no further than −106. In addition, species sequence comparisons (Martin *et al.*, 1988) (see Figure 5) show that this region is hypervariable when compared to the region −98 to −53 (our *Eco*RI site). In contrast the deletion of either the left- or right-hand part of this latter region often reduced, but did not abolish, *Sgs-3* expression, suggesting the presence of multiple sequences sufficient for developmentally specific expression. Similarly, Riddihough and Pelham (1987) described a 23-bp fragment from the *hsp27* promoter

which conferred ecdysone inducibility on an *hsp70–CAT* fusion gene transfected into *Drosophila* cells, but found that its disruption did not entirely eliminate ecdysone inducibility and were forced to conclude that further hormone responsive elements were present in the *hsp27* promoter. Jongens *et al.* (1988) came to similar conclusions with respect to the redundancy of tissue-specific elements in the analysis of the remote region containing the *Sgs-4* enhancer. By use of a transient expression assay they distinguished three regions which in pairwise combinations are sufficient to direct salivary gland expression, although none of the three is sufficient on its own. In this context, while our minimal construct $\Delta U_1 \Delta U_3 RH_{1-3}$ shows that the fusion gene may be expressed specifically in third instar glands with a 22-bp *Sgs-3* proximal upstream fragment (−98 to −77), we do not conclude that this alone contains all of the proximal regulatory element(s), as the 18-bp fragment U_3 (−76 to −59) is also sufficient for tissue-specific expression. It remains an open question as to whether these sequences are responsible for the tissue-specific expression or serve to stabilize the binding of factors to sequences either closer to the TATA sequence, e.g. straddling the *Eco*RI site from −58 to −50 conserved in both constructs (see below), or indeed in the distal (−600) region.

An important but unexplained result is the cell-to-cell variation in expression seen in a large proportion of lines carrying modified upstream sequences (Figure 2h, Table II). One possibility is that the modifications lower the probability either of the formation of a pre-activation complex earlier in development (e.g. a cell-specific chromatin structure) or of the efficiency of a protein–DNA complex during the period of expression. In either case expression will be subject to a stochastic process with regard to attaining the threshold for detection. Hiromi and Gehring (1987) also noted patchy colouration in embyros of UPHZ50H lines containing a *ftz/hsp70/lacZ* fusion gene. As in these stably transformed lines the DNA sequences are present in all cells, it suggests that we should reconsider the interpretation of mosaicism of expression in transient assays as a reflection of DNA distribution following injection (e.g. Jongens *et al.*, 1988).

The existence of multiple regulatory elements in eukaryotic enhancers and promoters now appears a general result (see Maniatis *et al.*, 1987, for review) and may reflect a regulation that is refined by the co-operative interaction of a number of transcription-factor binding motifs. Such interactions may not be limited to motifs within a physically contiguous regulatory element and complexes may be formed between distal and proximal regulatory sequences by DNA–protein and protein–protein interactions, with DNA bending or looping (see, for example, Schaffner *et al.*, 1988). In the case of *Sgs-3* the relative contribution of the proximal and distal elements and their interactions have still to be determined. However, the 63-bp deletion at −600 (Ramain *et al.*, 1988) and the ΔU_{1-3} deletion (the present study) both dramatically reduce expression of the *Sgs-3* promoter, while larger deletions either further upstream (Giangrande *et al.*, 1987) or between these elements (Ramain *et al.*, 1988) have little or no effect on transcript levels. In a few exceptional larvae of ΔU_{1-3} lines there is weak, cell-specific expression (see Results), which suggests that the distal element does not merely quantitatively enhance transcript levels but may contain tissue-specific elements, consistent with the observed cell- and stage-specific chromatin changes in this region (Ramain *et al.*, 1988).

Several lines of evidence show that ecdysone is involved in the regulation of the 68C glue gene cluster (see Meyerowitz *et al.*, 1987, for references). Whilst its role in *Sgs-3* induction appears somewhat indirect, and probably involves products of other loci, the hormone represses *Sgs-3* expression within minutes, which implies a direct interaction of the hormone–receptor complex with 68C sequences. Such an interpretation is supported by our results with the point mutation at −53 which, fortuitously, eliminates the developmental repression of the fusion gene and yet is apparently without effect upon induction. The restoration of repression in constructs carrying the H_3 sequence is at first sight surprising, but an inspection of the sequence (Figure 1) shows two tandem copies of the sequence GAATTG on the lower strand (+10 to −1), which is precisely the sequence we mutated at −53 to obtain our *Eco*RI site. Although indirect it is reasonable to suggest that this motif, or a part of it, is involved in the repression of *Sgs-3*.

What is the mechanism of hormonal repression? As repression is eliminated by a point mutation one possibility is that of competitive binding on DNA sequences in the proximal promoter, that interferes with the binding or activity of factors necessary for expression, e.g. the glucocorticoid inhibition of the human glycoprotein hormone α-subunit gene (Akerblom *et al.*, 1988). For *Sgs-3*, two situations are possible. In the first, two-site model, repression results from the binding of the hormone–receptor complex in the region of the *Eco*RI site whilst induction is mediated by the hormone-induced products of other loci (possibly interacting with the hormone–receptor complex itself) which interact with other *Sgs-3* sequences. In the second, a single site is modified by stage-specific factors bound to sequences flanking the *Eco*RI site, thus creating a site that mediates the induction of *Sgs-3* by an interaction with the hormone–receptor complex (and/or hormone-induced products) that is distinct from that which represses *Sgs-3* expression at the end of the instar. The results with the H_3-containing constructs favour the former as it appears that an element in the fusion gene leader creates a novel site that mediates repression. It may be possible to distinguish between these models by binding studies using *Drosophila* receptor proteins, currently being characterized in a number of laboratories, with both wild-type and mutant *Sgs-3* templates.

Materials and methods

Construction of transposons

The construction of the glAd transposon is given in detail in Bourouis and Richards (1985b). The *Hin*dIII linkers used for the fusion are at +10 on the *Sgs-3* leader and at +13 on the *Adh* larval leader [and not +9 as reported by Bonner *et al.* (1984)—see Figure 1]. Mutants were constructed and sequenced in pEMBL182 (C.Mettling, unpublished data) which contains a *Xho*I site in the polylinker. The *Eco*RI and *Hin*dIII sites of this polylinker were destroyed for the purposes of later constructs (see below). For glAd130 and glAd130I the *Sal*I (−130)–*Xho*I fragment was inserted in the *Sal*I site of Carnegie 20 (Rubin and Spradling, 1983), for all other constructs the equivalent fragment was inserted in the *Sal*I site of glΔS (Bourouis and Richards, 1985b) which contain 2.6 kb of *Sgs-3* 5′ flanking sequence (see Figure 1). In the latter we selected and analysed only those constructs in which the natural orientation of 5′ flanking sequences is restored.

A point mutation was introduced at −53 (G→C), so as to create an *Eco*RI site, using the double-primer approach (Norris *et al.*, 1983) modified by filtering the elongation reaction through nitrocellulose filters to enrich for double-stranded DNA as suggested by Nakamaye and Eckstein (1986) for

a related oligonucleotide mutagenesis technique. The resulting glAdE pEMBL subclone was digested with *Eco*RI, filled in and a *Sal*I linker added. Digestion by *Sal*I and religation resulted in pEMBLΔU_{1-3}. For the remaining constructs the glAdE subclone was digested by *Sal*I + *Eco*RI or *Eco*RI + *Hin*dIII and sets of overlapping complementary oligonucleotides were inserted by shotgun ligation (Grundström *et al.*, 1985) to give the derivatives shown in Figure 1. In the H_1 box the *hsp70* sequences (Ingolia *et al.*, 1980) extend to −50, thus avoiding the presence of a heat shock response element in the constructs.

Isolation and characterization of transformed strains

Transposons (300 μg/ml) were co-injected with the integration-defective helper plasmid pπ25.7WC (300 μg/ml) (Karess and Rubin, 1984) into embryos of the *b* Adh^{nLA248}; ry^{506} strain, constructed by M.Bourouis from *w*, *b* Adh^{nLA248} (supplied by W.Moses, Cambridge) and ry^{506} (W.Bender, Harvard). This recipient strain is less sensitive to infections than the Adh^{n4} strain used previously for glAd lines (glAd:1 to glAd:4) (Bourouis and Richards, 1985b). Homozygous lines carrying single insertions were derived by crosses to appropriate balancers. For ADH staining, second chromosome hemizygous viable insertions were examined in the parental line selected for the ry^{+} phenotype for several generations; third chromosome lines were further selected for the *b* Adh^{nLA248} chromosome following the balancer crosses that assigned chromosome location. All lines were examined by Southern analysis to verify the presence of a single non-rearranged insertion of the transposon (data not shown).

The ADH histochemical assay

Animals (in groups of five) were dissected on ice in phosphate-buffered saline and stained essentially as described in Martin *et al.* (1986) with modifications suggested by Dr J.Deutsch (Institut Jacques Monod, Paris), i.e. the substitution of a mixture of ethanol and iso-propanol (2 vol:1 vol) for 2-butanol as substrate and fixation in 45% acetic acid after staining. Each animal was classed on a four-point scale (−, +, ++, +++ for no, weak, intermediate and strong staining) and the presence or absence of mosaicism noted (see Figure 2h and Tables I and II). Lines showing such patchy expression were retested several generations later to confirm that it was not a consequence of growth conditions in a given vial. Exceptionally, we observed staining of cells in the proventriculus, similar to that described for larvae transformed with an *Sgs-4/Adh* construct (tGARP301 lines) (Shermoen *et al.*, 1987).

RNA extraction and primer extension analysis of transcripts

RNA was extracted from third instar larvae and prepupae using a guanidinium/LiCl based procedure. For 6-h prepupae, animals were collected as white prepupae (+/− 15 min) and held on moist filter paper for 6 h at 25°C. Pools of 10−15 animals were homogenized in 150 μl of a solution of 5 M guanidinium monothiocyanate, 50 mM Tris−HCl pH 7.6, 10 mM EDTA, 2% Sarkosyl, 10 mM ribonucleoside−vanadyl complex, 5% β-mercaptoethanol. Then 150 μl of 8 M LiCl was added, the tube vortexed and left on ice overnight. The precipitate was recovered by centrifugation, resuspended in 100 μl of 0.1 M NaOAc, 0.1% SDS and extracted successively with phenol, phenol−chloroform and chloroform. The final supernatant was reprecipitated with 100 μl of 8 M LiCl for >3 h on ice at 4°C. The precipitate was recovered by centrifugation, resuspended in 100 μl of distilled water, adjusted to 0.3 M NaOAc pH 4.5 and reprecipitated with 2.5 vol of ethanol.

For primer extension, 5 μg of total RNA were co-precipitated with 0.1−0.5 ng of the appropriate ^{32}P-labelled oligonucleotide. Samples were then processed as described by Schmitt *et al.* (1987), and the products analysed on a denaturing 10% polyacrylamide gel. For RNA sequencing, 10 μg of total RNA was used for each base reaction as described by Schmitt *et al.* (1987) using a nucleotide ratio of 0.4:1 (ddNTP:dNTP). Primers were as follows:

Adh primer A, $^{5'}{}_{(+53)}$TTTAGCAGGCTCTTTCGATCTGTT$^{3'}{}_{(+30)}$
primer B, $^{5'}{}_{(+35)}$TCTGTTTCAATTGGAAGAGG$^{3'}{}_{(+16)}$
Sgs-3, $^{5'}{}_{(+54)}$GGGCGGTAGCAATGGTC$^{3'}{}_{(+38)}$

Acknowledgements

We thank Professor P.Chambon for his continued support and interest in this work. We thank Frank Ruffenach and Adrien Staub for the synthesis of oligonucleotides. We are grateful to Hinrich Gronemeyer and Bohdan Wasylyk for a critical reading of a draft of this manuscript which was prepared with the help of B.Boulay. This work was supported by grants from the CNRS, the INSERM, the Fondation pour la Recherche Médicale, the Ministère de la Recherche et de l'Enseignement Supérieur and the Association pour la Recherche sur le Cancer. M.M. is a research student funded by the Ministère de la Recherche et de la Technologie and the Université Louis Pasteur, Strasbourg. A.G. is funded by the Consiglio Nazionale delle Ricerche Progetto Finalizzato 'Ingegneria Genetica' U.O. 47 No. 8686/51, Istituto di Genetica, Università di Bari.

References

Akerblom,I.E., Slater,E.P., Beato,M., Baxter,J.D. and Mellon,P.L. (1988) *Science*, **241**, 350−353.
Berendes,H.D. and Ashburner,M. (1978) The salivary glands. In Ashburner,M. and Wright,T.R.F. (eds), *The Genetics and Biology of Drosophila*. Academic Press, London and New York, Vol. 2B, pp. 453−498.
Bienz,M., Saari,G., Tremml,G., Müller,J., Züst,B. and Lawrence,P.A. (1988) *Cell*, **53**, 567−576.
Bonner,J.J., Parks,C., Parker-Thornburg,J., Mortin,M.A. and Pelham,H.R.B. (1984) *Cell*, **37**, 979−991.
Bourouis,M. and Richards,G. (1985a) *Cell*, **40**, 349−357.
Bourouis,M. and Richards,G. (1985b) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 355−360.
Bray,S.J., Johnson,W.A., Hirsh,J., Heberlein,W. and Tjian,R. (1988) *EMBO J.*, **7**, 177−188.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349−383.
Bucher,P. and Trifonov,E.N. (1986) *Nucleic Acids Res.*, **14**, 10009−10026.
Chia,W., Savakis,C., Karp,R., Pelham,H. and Ashburner,M. (1985) *J. Mol. Biol.*, **186**, 679−688.
Fischer,J.A. and Maniatis,T. (1988) *Cell*, **53**, 451−461.
Garabedian,M.J., Shepherd,B.M. and Wensink,P.C. (1986) *Cell*, **45**, 859−867.
Garfinkel,M.D., Pruitt,R.E. and Meyerowitz,E.M. (1983) *J. Mol. Biol.*, **168**, 765−789.
Giangrande,A., Mettling,C. and Richards,G. (1987) *EMBO J.*, **6**, 3079−3084.
Grundström,T., Zenke,M.W., Wintzerith,M., Matthes,H.W.D., Staub,A. and Chambon,P. (1985) *Nucleic Acids Res.*, **13**, 3305−3316.
Hiromi,Y. and Gehring,W.J. (1987) *Cell*, **50**, 963−974.
Hiromi,Y., Kuroiwa,A. and Gehring,W.J. (1985) *Cell*, **43**, 603−613.
Hultmark,D., Klemenz,R. and Gehring,W.J. (1986) *Cell*, **44**, 429−438.
Ingolia,T.D., Craig,E.A. and McCarthy,B.J. (1980) *Cell*, **21**, 669−679.
Jongens,T.A., Fowler,T., Shermoen,A.W. and Beckendorf,S.K. (1988) *EMBO J.*, **7**, 2559−2567.
Karess,R.E. and Rubin,G.M. (1984) *Cell*, **38**, 135−146.
Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) *Science*, **236**, 1237−1245.
Martin,C.H., Mayeda,C.A. and Meyerowitz,E.M. (1988) *J. Mol. Biol.*, **201**, 273−287.
Martin,P., Martin,A., Osmani,A. and Sofer,W. (1986) *Dev. Biol.*, **117**, 574−580.
Meyerowitz,E.M., Vijay Raghavan,K., Mathers,P.H. and Roark,M. (1987) *TIG*, **3**, 288−293.
Nakamaye,K.L. and Eckstein,F. (1986) *Nucleic Acids Res.*, **14**, 9679−9698.
Norris,K., Norris,F., Christiansen,L. and Fill,N. (1983) *Nucleic Acids Res.*, **11**, 5103−5112.
Ramain,P., Giangrande,A., Richards,G. and Bellard,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2718−2722.
Riddihough,G. and Pelham,H.R.B. (1987) *EMBO J.*, **6**, 3729−3734.
Romano,C.P., Bienz-Tadmor,B., Mariani,B.D. and Kafatos,F.C. (1988) *EMBO J.*, **7**, 783−790.
Rubin,G.M. and Spradling,A.C. (1982) *Science*, **218**, 348−353.
Rubin,G.M. and Spradling,A.C. (1983) *Nucleic Acids Res.*, **11**, 6341−6351.
Schaffner,G., Schirm,S., Müller-Baden,B., Weber,F. and Schaffner,W. (1988) *J. Mol. Biol.*, **201**, 81−90.
Schmitt,P., Gattoni,R., Keohavong,P. and Stevenin,J. (1987) *Cell*, **50**, 31−39.
Shermoen,A., Jongens,J., Barnett,S., Flynn,K. and Beckendorf,S.K. (1987) *EMBO J.*, **6**, 207−214.
Simon,M.C., Fisch,T.M., Benecke,B.J., Nevins,J.R. and Heintz,N. (1988) *Cell*, **52**, 723−729.
Vijay Raghavan,K., Crosby,M.A., Mathers,P.H. and Meyerowitz,E.M. (1986) *EMBO J.*, **5**, 3321−3326.
Xiao,H. and Lis,J.T. (1988) *Science*, **239**, 1139−1142.

Received on October 10, 1988; revised on December 2, 1988